

Advanced Ensemble Learning with Pruning Techniques for Detecting DDOS Attacks in IOT Networks

¹M. Mutharasu, ²Meghana Chowdary. P, ³Mizba Kousar. S

¹Assistant Professor, C.S.E (Cyber Security), Madanapalle Institute of Technology & Science, Madanapalle, India ^{2,3}UG Scholar, C.S.E (Cyber Security), Madanapalle Institute of Technology & Science, Madanapalle, India E-mail: <u>¹mutharasum@mits.ac.in, ²potturimeghanachowdary@gmail.com</u>, <u>3kousarmizba3@gmail.com</u>

Abstract - IoT networks are vulnerable to Distributed Denial of Service (DDoS) attacks, which can cause network instability and compromised data integrity. In order to improve detection efficiency and accuracy, this research suggests a revolutionary deep ensemble learning architecture that incorporates pruning strategies. The system uses a Voting Classifier that combines K-Nearest Neighbours, Decision Trees, and Logistic Regression, and a Stacking Classifier that combines Random Forest, Gradient Boosting, and Naïve Bayes models. The machine learning pipeline is automated and optimized by a TPOT Classifier, and redundant models are eliminated for computational efficiency through pruning. The system efficiently differentiates between typical and malicious traffic patterns by using an extensive dataset of network flow characteristics, including packet lengths, inter-arrival periods, and flag counts. The model's exceptional performance is demonstrated by evaluation criteria such as accuracy, precision, and recall, which show that it achieved over 98% detection accuracy with less false positives. IoT network vulnerabilities are addressed by this resource-efficient and scalable method, which provides a strong real-time threat detection solution.

Keywords: DDoS attack detection, IoT security, deep ensemble learning, pruning techniques, Stacking Classifier, Voting Classifier, TPOT Classifier, network flow metrics, machine learning.

I. INTRODUCTION

The Internet of Things has becomegame-changing technology that is enabling automation and innovation in variety of industries, including healthcare, and transportation, with billions of devices connected globally. However, because linked gadgets spread quickly, there are already significant cybersecurity problems. Distributed Denial of Service (DDoS) attacks are among the most prevalent and destructive of these problems. These assaults exploit the interconnection of IoT devices by overloading networks with malicious traffic, leading to data breaches, outages, and financial losses. Because of the inherent limitations of IoT devices, including their limited computational power, storage, and bandwidth, it is more challenging to identify and mitigate such attacks effectively.

In IoT environments, conventional DDoS detection techniques, such as signature-based and threshold-based methods, are frequently insufficient. Signature-based techniques have trouble spotting novel or changing threats since they depend on pre-established patterns of recognized assaults. Contrarily, threshold-based systems impose preset restrictions on network traffic, which leaves them vulnerable to high false-positive rates and a lack of flexibility in response to changing network conditions. These restrictions show how urgently creative solutions that are adapted to the particulars of IoT networks are needed.

To improve the precision, scalability, and computational efficiency of DDoS attack detection in IoT networks, this study introduces a unique framework that blends deep ensemble learning approaches with pruning methods. In order to capitalize on each machine learning model's unique capabilities while minimizing its shortcomings, the suggested solution uses a hybrid ensemble method. In particular, the framework integrates Stacking Classifier, which combines techniques like Random Forest, Gradient Boosting, and Naïve Bayes. A meta-classifier aggregates the outputs of basic models to enhance decision-making. Voting Classifier, which uses K-Nearest Neighbours (KNN), Decision Trees (DT), and Logistic Regression (LR) to reach a consensus through best-performing majority voting. The models and hyperparameters for the detection job are found by the automated machine learning pipeline optimizer TPOT Classifier.

One important feature is the addition of pruning algorithms to the ensemble architecture. Pruning lowers computational overhead and improves system efficiency by removing unnecessary and redundant models from the ensemble. In IoT contexts with limited resources, where reducing computing demands is crucial, this optimization is very beneficial. The suggested detection algorithm is based on an extensive dataset that has been enhanced with network flow



https://doi.org/10.47001/IRJIET/2025.INSPIRE34

International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

characteristics. Features including flow duration, packet lengths, inter-arrival times, and flag counts (e.g., SYN, FIN, ACK) offer a solid basis for distinguishing between benign and malicious traffic patterns. Even in real-time situations, the system may successfully identify anomalies suggestive of DDoS attacks by examining these traits.

The proposed framework is evaluated using key performance metrics, including as accuracy, precision, recall, and F1-score, to ensure that it is reliable and strong. Experimental results show that, in comparison to traditional machine learning methods, the deep ensemble model achieves higher detection accuracy while maintaining a low falsepositive rate. These results highlight how well the system can handle the crucial issues of efficiency and scalability in IoT security.

II. RELATED WORK

1. Machine Learning-Based DDoS Detection in IoT Networks

Machine learning (ML) approaches for identifying DDoS assaults in IoT networks have been studied recently. Malicious traffic patterns can be successfully classified using ML-based techniques, such as supervised and unsupervised learning models. For anomaly detection, methods such as deep learning models, support vector machines (SVM), and decision trees have been used. In order to spot questionable activity, these models examine network traffic characteristics including packet flow rates and protocol types. Even if ML-based systems increase the accuracy of detection, managing massive IoT traffic and modifying models to account for changing attack patterns continue to be difficult tasks.

2. Ensemble Learning Approaches for Cybersecurity

Ensemble learning techniques use several base classifiers to increase detection accuracy, they have become more and more popular in the cybersecurity field. To increase the resilience of intrusion detection systems, strategies like bagging, boosting, and stacking classifiers have been used. To enhance classification performance, researchers have combined classifiers like Random Forest, Gradient Boosting, and Naïve Bayes. Compared to single classifiers, ensemble models lower false positives and offer superior generalization capabilities. Ensemble models' computational overhead is still a problem, though, and calls for optimization techniques like pruning and attack patterns. However, because harmless changes in network traffic might be mistaken for hostile activity, these techniques frequently result in significant false positive rates. In order to get over these obstacles, hybrid detection methods that combine several strategies have become more popular recently.

3. Pruning Techniques for Optimized Machine Learning

Pruning is a model optimization approach that improves computational efficiency by removing unnecessary or insignificant features and classifiers. Pruning lowers resource usage in IoT-based DDoS detection systems in cybersecurity. Pruning has been used by researchers to reduce complexity while preserving detection accuracy in ensemble and deep learning models. To make detection pipelines more efficient, methods like feature selection, weight pruning, and structured pruning have been investigated. Pruning ensures minimal latency and less processing power while enabling real-time attack detection in IoT security solutions.

4. Deep Learning Models for DDoS Detection

Deep learning has transformed cybersecurity by offering scalable and incredibly accurate DDoS assault detection tools. Recurrent neural networks and convolutional neural networks have been used to evaluate network traffic aspects and accurately identify attack patterns. To improve real-time anomaly detection, researchers have created hybrid deep learning models that combine CNNs with LSTM networks. Deep learning models are successful, but their application in resource-constrained IoT environments is difficult due to their high processing requirements.

5. Automated Machine Learning (AutoML) for Intrusion Detection

To automate the model selection and optimization procedure, Automated Machine Learning (AutoML) solutions like TPOT have been established. TPOT selects the best effective models for a given dataset by use genetic programming to investigate several ML processes. Because auto ML eliminates the need for manual feature engineering and hyperparameter tuning, it improves the development of DDoS detection systems. TPOT-based detection frameworks have been used by researchers, and they perform better than conventional ML techniques. Regular updates are necessary for AutoML-generated models to stay abreast of evolving cyberthreats.

6. Feature Engineering for DDoS Attack Identification

Enhancing the capabilities of DDoS detection algorithms requires feature engineering. Research has focused on finding and eliminating the most relevant network traffic characteristics in order to distinguish between malicious and genuine traffic. Using techniques like SelectKBest and Recursive Feature Elimination, critical features such as flow



Volume 9, Special Issue INSPIRE'25, pp 210-216, April-2025

https://doi.org/10.47001/IRJIET/2025.INSPIRE34

International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

time, packet lengths, and inter-arrival intervals have been found. Efficient feature selection enhances the interpretability and efficacy of machine learning models, enabling more rapid and accurate attack detection in Internet of Things networks.

7. Real-Time DDoS Detection in IoT Networks

It is still very challenging to identify DDoS attacks in real time in IoT networks due to the vast volume of data and the low processing capability of IoT devices. Researchers have suggested deep learning and lightweight machine learning models that are capable of real-time intrusion detection.In order to reduce latency and increase detection speed, federated learning and edge computing techniques have been investigated. Adaptive thresholding and incremental learning are used by real-time detection systems to dynamically detect anomalies. Research is still being done to provide real-time processing without sacrificing accuracy.

III. PROPOSED WORK

Using an automated TPOT classifier in conjunction with stacking and voting classifiers, a deep ensemble learning technique is used to detect DDoS attacks in IoT networks. The model increases detection accuracy while decreasing false positives by utilizing RF, Gradient Boosting, Logistic Regression, Decision Tree, KNN, NaïveBayes, and AdaBoost..Pruning approaches remove duplicate classifiers to maximize computing performance. This scalable system guarantees real-time threat detection, tackling IoT security issues with little resource overhead and great precision.

IV. METHODOLOGY

A. Data Collection

Network traffic data is gathered from publicly accessible datasets, simulated attack settings, or real-time network records in order to detect DDoS attacks in IoT networks. Eighty network flow variables, including flow duration, packet lengths, inter-arrival durations, flag counts (SYN, FIN, ACK), and the total number of forward and backward packets, are included in the dataset utilized in this study. These characteristics aid in distinguishing between benign and malevolent traffic patterns. To improve model performance, data preprocessing methods such as feature selection, normalization, and encoding are used. Machine learning algorithms employ label encoding to translate categorical data into numerical values. Furthermore, to comprehend data distribution, spot abnormalities, and eliminate superfluous features, exploratory data analysis, or EDA, is carried out. The performance of various classifiers is then assessed by dividing the dataset into training and testing subsets. In resourceconstrained IoT contexts, this thorough data collection method guarantees precise and effective DDoS detection.

B. Preparation of Dataset

Data preprocessing, feature selection, and data splitting are some of the stages involved in preparing the dataset for DDoS attack detection in IoT networks. First, raw network traffic data is gathered, including flow duration, packet sizes, inter-arrival periods, and flag counts, among other features. Duplicate records are eliminated and missing values are handled using imputation techniques to guarantee correctness and consistency. To find the most pertinent features, feature selection is carried out using the SelectKBest approach, which lowers dimensionality and boosts model performance. Label encoding is used to translate categorical features, like protocol kinds, into numerical values. The performance of machine learning models is then improved by normalizing the dataset to guarantee consistent feature scaling.Exploitary Data Analysis (EDA) is used to identify abnormalities and display the distribution of data. To provide efficient model evaluation and generalization for real-time DDoS detection in IoT contexts, the dataset is finally divided into training and testing subsets.

C. Preprocessing and Feature Extraction

Enhancing the precision and effectiveness of DDoS attack detection in IoT networks requires preprocessing and feature extraction. To preserve data integrity, the preparation step starts with deleting duplicate or unnecessary entries and managing missing values using imputation techniques. In order to eliminate biases in machine learning models, feature scaling approaches like normalization and standardization are used to guarantee consistency in numerical data. In order to make categorical features—such as protocol types compatible with classifiers, label encoding is utilized to translate them into numerical representations.

Finding the most relevant network traffic characteristics that distinguish between benign and malicious activity is the aim of feature extraction. Based on statistical scores, the bestperforming features are found using the Select K Best approach. Flow duration, packet inter-arrival periods, flag counts (SYN, FIN, ACK), and the total number of forward and backward packets are important extracted features. Model accuracy and computational efficiency are improved through feature extraction, which lowers dimensionality and gets rid of redundant features. The dataset is ideal for deep ensemble learning-based DDoS detection in IoT networks thanks to these preprocessing and extraction procedures.



https://doi.org/10.47001/IRJIET/2025.INSPIRE34

International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

D. Selection of Model and Training

Choosing a model and training are essential steps in creating a DDoS attack detection system that works for IoT networks. To improve detection accuracy while reducing computing overhead, this study uses a deep ensemble learning strategy that integrates a Stacking Classifier, Voting Classifier, and TPOT Classifier. Utilizing the advantages of several methods, the Stacking Classifier uses Random Forest and Gradient Boosting as basis models and Logistic Regression as the final estimator. By combining predictions from AdaBoost, KNN, Decision Trees, Naïve Bayes, and Logistic Regression, the Voting Classifier ensures robustness by majority voting. The TPOT Classifier uses genetic programming to optimize automating performance by model selection and hyperparameter adjustment.

In order to effectively distinguish between malicious and legitimate traffic, training entails extracting features from network flow data, such as packet statistics, inter-arrival periods, and flag counts. By removing unnecessary elements, pruning strategies enhance the model's accuracy and computing efficiency. Metrics including accuracy, precision, recall, and F1-score are used to evaluate performance, and the results show that the suggested ensemble model greatly improves DDoS detection, making it a scalable and effective IoT security solution.

E. Threat Classification and Detection

By examining network traffic patterns and anomalies, threat classification and detection in IoT networks entails spotting hostile activity, especially DDoS attacks. To increase accuracy, this study uses deep ensemble learning andintegrates TPOT, Voting, and Stacking classifiers. To differentiate between threats and regular traffic, the classification process makes use of important network flow characteristics such packet lengths, inter-arrival durations, and flag counts. The system improves computing efficiency by employing pruning approaches in conjunction with machine learning models such as RF, GB, and AdaBoost. Accuracy and F1-score are two performance indicators that confirm the model's efficacy in real-time DDoS detection, guaranteeing strong IoT network security.

F. Evaluation

A dataset with important network flow properties is used to test the deep ensemble learning method, which includes Stacking, Voting, and TPOT classifiers. The robustness of the method is highlighted by experimental findings that show that the TPOT and Naïve Bayes classifiers attain 98% accuracy, while the Stacking and Voting classifiers obtain 99%. By removing unnecessary models and cutting processing overhead without sacrificing accuracy, pruning approaches further improve computational efficiency. By successfully reducing false positives, the model guarantees accurate realtime detection of harmful activity. The efficacy of our method in managing high-volume IoT traffic is confirmed by comparison with conventional machine learning models. All things considered, the assessment confirms that the suggested method is a scalable and effective way to improve IoT network security against changing DDoS attacks.





V. RESULTS AND DISCUSSION

The suggested deep ensemble learning-based DDoS detection system's outcomes show how well it can detect malevolent attacks in Internet of Things networks. The robustness of the ensemble approach is validated by experimental evaluations, which demonstrate that the TPOT and Naïve Bayes classifiers obtain 98% detection accuracy, while the Stacking and Voting classifiers achieve a high detection accuracy of 99%. The system's capacity to discriminate between legitimate and malicious traffic is greatly improved by the incorporation of feature extraction techniques, which include network flow metrics like packet lengths, inter-arrival periods, and flag counts. By removing redundant models, pruning strategies further maximize computational efficiency and guarantee real-time detection with no overhead.

The system successfully lowers false positives and false negatives, according to evaluation criteria including precision, recall, and F1-score, making it a dependable IoT security solution. The suggested solution addresses the dynamic nature of cyber threats in IoT contexts and shows better scalability and adaptability than classic machine learning models. These findings demonstrate that deep ensemble learning may be used to improve threat detection, guaranteeing safe and robust IoT systems while reducing computational complexity.



International Research Journal of Innovations in Engineering and Technology (IRJIET) ISSN (online): 2581-3048

Volume 9, Special Issue INSPIRE'25, pp 210-216, April-2025

https://doi.org/10.47001/IRJIET/2025.INSPIRE34

International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)



Fig 2: Confusion Matrix of STACKING, TPOT, VOTING CLASSIFIERS



Fig 3: Confusion Matrix of NAÏVEBAYES, ADABOOST CLASSIFIERS

VI. CONCLUSION

The suggested deep ensemble learning-based DDoS detection system for IOT networks successfully tackles the increasing threats to cybersecurity. The system maintains computational efficiency while achieving excellent detection accuracy by combining pruning approaches with TPOT, Voting, and Stacking classifiers. By utilizing the advantages of several algorithms, the integration of various models—including RF, GB, Logistic Regression, Decision Tree, KNN, Naïve Bayes, and AdaBoost—ensures reliable threat categorization.



Fig 4: Performance metrics

By improving model selection and hyperparameter tuning, TPOT's automated pipeline optimization makes the system more responsive to changing attack patterns. In order to differentiate between DDoS assaults and regular traffic, it is essential to extract features from network flow data, such as packet lengths, inter-arrival periods, and flag counts. By removing redundant models, pruning strategies lower resource consumption and enable real-time detection that is appropriate for IoT applications with limited resources.



Fig 5: BAR GRAPH

Experiments show that the system performs better than conventional machine learning techniques, with detection accuracies of 98% for TPOT and Naïve Bayes classifiers and 99% for Stacking and Voting classifiers. Furthermore, the incorporation of performance measurements like precision, recall, and F1-score guarantees thorough evaluation of model's efficacy, showcasing its capacity to reduce false positives and false negatives. The system's scalability allows it to be used in a variety of IoT infrastructures, such as healthcare systems, and industrial IoT settings. The suggested ensemble learning approach improves resilience and adaptability in contrast to traditional signature-based or threshold-based detection mechanisms, which find it difficult to adjust to changing attack patterns. Additionally, the system's capacity to automate feature selection and optimization speeds up deployment and minimizes manual labor.

Adaptive learning techniques that dynamically update the model in response to novel attack patterns can be included into future improvements to further increase the system's resilience. Adding user behavior analytics and applicationlayer features to the feature set can improve classification accuracy and offer more in-depth understanding of network traffic irregularities. The detection system can be made even more effective in extensive IoT deployments by connecting it with edge computing frameworks, which can lower latency and enhance real-time threat mitigation. Cross-platform testing





Volume 9, Special Issue INSPIRE'25, pp 210-216, April-2025 https://doi.org/10.47001/IRJIET/2025.INSPIRE34

International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

in various IoT contexts will confirm the model's applicability and generalizability to various network circumstances.

A more complete security solution may also be offered by expanding the system's detection capabilities to include additional cyberthreats like SQL injection and Man-in-the-Middle (MitM) assaults. By adding real-world data, a user feedback mechanism can improve detection algorithms, further lowering false positives and increasing overall reliability. In order to guarantee that the system complies with industry best practices and promotes broad adoption, cooperation with IoT security standards organizations is necessary.

All things considered, the suggested deep ensemble learning-based method offers a scalable, precise, and computationally effective solution for DDoS attack detection, marking a substantial leap in IoT security. Its potential to protect vital IoT infrastructures is highlighted by its capacity to improve network resilience and reduce cyber threats. The system provides a useful and efficient method for real-time threat identification by utilizing a variety of machine learning models, maximizing computing efficiency through pruning, and integrating automation with TPOT. This research makes a addition to cybersecurity since the need for intelligent and adaptable security solutions will only grow as IoT networks continue to grow.

REFERENCES

- Ahmad, S., Bakar, A. H. A., & Ali, M. I. (2021). Hybrid Model for DDoS Attack Detection in IoT Using Machine Learning Techniques. Future Generation Computer Systems, 119, 41–54.
- [2] Ali, W. M. M., Abed, H. S., & Zawawi, N. A. A. (2022). An Efficient DDoS Detection System for IoT Using Deep Learning Techniques. IEEE Access, 10, 20645–20659.
- [3] Al-Mamun, M., Abdullah, A. I., & Khan, M. K. (2021). Ensemble Learning for DDoS Attack Detection in IoT Networks. Journal of Information Securityand Applications, 57, 102758.
- [4] Cheng, Y., Wei, J., & Li, H. (2021). IoT Security and DDoS Attack Prevention Using Federated Learning. Computers & Security, 110, 102460.
- [5] Feng, C., Liu, M., & Wang, X. (2020). A Hybrid Deep Learning Model for IoT Intrusion Detection. Journal of Network and Computer Applications, 112, 102888.
- [6] F. A. Qureshi, H. H. Alazab, and A. J. H. Alzahrani, "DDoS Attack Detection in IoT Networks: A Review," Computers & Security, vol. 110, pp. 102427, 2021.
- [7] Gupta, K. K., Shafique, H., Awan, M. S., & Babar, M. U. (2021). A Survey on DDoS Attack Detection

Mechanisms in IoT Networks. IEEE Access, 9, 102988–103005.

- [8] Hussain, S. A., Raza, A., & Khan, N. (2021). Adaptive DDoS Attack Detection in IoT Environments Using AI-Based Techniques. Sensors, 21(5), 1542.
- [9] Islam, M. A., Ahmad, S., & Bakar, H. (2022). A Novel Approach for Real-Time DDoS Detection in IoT Networks Using Ensemble Learning. Future Internet, 14(3), 45.
- [10] Javed, T., Iqbal, S., & Khalid, A. (2020). Anomaly-Based DDoS Attack Detection Using AI-Driven Techniques. Computers & Security, 109, 102457.
- [11] Kumar, A., Sharma, P., & Gupta, R. (2021). Enhancing IoT Network Security Through Federated Learning-Based DDoS Attack Detection. IEEE Transactions on Industrial Informatics, 17(9), 6543-6555.
- [12] Lin, J., Deng, R., & Chen, Y. (2021). IoT Security Threats and Countermeasures: A Machine Learning Perspective. ACM Computing Surveys, 54(2), 1–36.
- [13] Mahmood, T., Khan, Z., & Usman, M. (2021).
 Lightweight Intrusion Detection System for IoT Using Feature Pruning and Machine Learning Models.
 Computers & Security, 112, 102888.
- [14] Patel, V. K., & Sharma, R. K. (2021). DDoS Detection in IoT Networks Using Stacking Classifier Approach. Future Internet, 13(10), 272.
- [15] Qureshi, F. A., Alazab, H. H., & Alzahrani, A. J. H. (2021). DDoS Attack Detection in IoT Networks: A Review. Computers & Security, 110, 102427.
- [16] Rahman, M. M., & Azim, M. (2021). Anomaly-Based Intrusion Detection System for IoT Using Deep Learning Techniques. IEEE Access, 9, 104567– 104580.
- [17] Santos, J. L. M. B., Silva, E. M. F. D., & Silva, R. D. S. (2020). An Ensemble Learning Approach for DDoS Attack Detection in IoT Networks. IEEE Latin America Transactions, 18(6), 1058–1065.
- [18] Sun, X., Guo, J., & He, Z. (2022). Improving DDoS Attack Detection with Hybrid Machine Learning Models. Future Generation Computer Systems, 130, 45–59.
- [19] Verma, R., Srivastava, A. K., & Gupta, R. K. (2022). Deep Learning-Based Detection of DDoS Attacks in IoT Networks. Journal of Computer Networks and Communications, 2022, 2039265.
- [20] Zhang, H., Yuan, Y., & Zhou, L. (2021). Adaptive Deep Learning Framework for IoT Cybersecurity. IEEE Transactions on Dependable and Secure Computing, 19(1), 128–140.



Citation of this Article:

M. Mutharasu, Meghana Chowdary. P, & Mizba Kousar. S. (2025). Advanced Ensemble Learning with Pruning Techniques for Detecting DDOS Attacks in IOT Networks. In proceeding of International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25), published by *IRJIET*, Volume 9, Special Issue of INSPIRE'25, pp 210-216. Article DOI https://doi.org/10.47001/IRJIET/2025.INSPIRE34
